International Journal of Exploring Emerging Trends in Engineering (IJEETE)
Vol. 03, Issue 02, Mar-Apr, 2016 Pg. 126 - 131 WWW.IJEETE.COM

REAL-TIME IMPLEMENTATION OF BINARY MASK ALGORITHM FOR HEARING PROSTHETICS

[1]Swati R. Pawar, [2]Hemantkumar B. Mali
[1, 2] Dept. of Electronics and Telecommunication Engg
Savitribai Phule University of Pune
SITS, Narhe, Pune-41
Pune, India.

ABSTRACT - *As the current speech enhancement algorithms can give results for improved speech audibility only. So, our main motivation is to develop speech enhancement algorithm that would improve performance in speech for hearing impaired persons. In this paper, an algorithm is presented for improving speech intelligibility. Various speech enhancement algorithms were developed but only some of them can be used for real time hearing aid applications. This proposed algorithm can be used for practical hearing prosthetic devices. Implementation of the binary masking algorithm uses a bank of band-pass filters to perform mapping of signals. Also, classification is performed with a signal-to-noise (SNR) estimate and a comparator. This includes spatial filtering method, classification of signals such as original and noisy signal. After this based on SNR threshold level signals are recombined to obtain reduced noise level in speech signal. In this, matlab implementation of binary mask algorithm is provided which shows better results for speech intelligibility as compared to other algorithms.*

Keywords— Binary Mask Algorithm (BMA), Signal-to-noise ratio (SNR), Speech intelligibility, Speech quality, Time-Frquency (T-F), Independent component analysis (ICA).

I. INTRODUCTION

Hearing impaired listeners find more difficulty in distin- guishing speech in noisy environments as compared to persons with normal hearing. While using correct hearing devices hear- ing impaired listeners find it difficult in interpreting speech. The existing methods used in available speech algorithms gives better results for quality of speech signal. While clarity of speech will give poor performance. Because of this hearing impaired listeners having problems in communication. This speech intelligibility problem is significantly removed by using binary mask algorithm. This proposed method is suitable for hearing aid application. Time frequency masking technique is introduced in this algorithm. Time frequency masks often take the values of 'zero' and 'one', resulting in mixture of two signals. Binary mask is calculated in classification stage and noise signal is removed. Second, methods used for calculating masks are based on computational auditory scene analysis [15]

II. COMPARISON OF PROPOSED M ETHOD WITH EXISTING METHODS

The existing methods used in hearing aids improve only speech quality of signals and ease of listening. But the speech intelligibility improvement for hearing aid application is limited. So, it is important to develop speech separation algorithms for hearing aid application that [15] will increase the intelligibility of the speech more effectively.

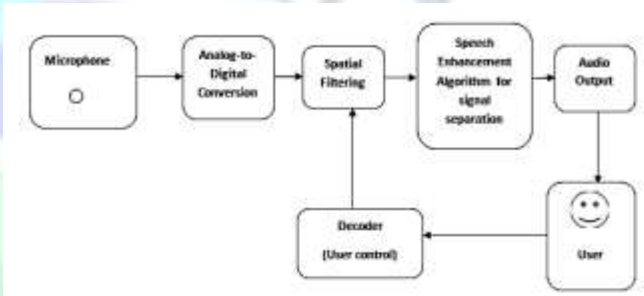

Fig. 1. Binary Mask Algorithm for enhancing speech intelligibility.

The basic technique is the time frequency masking for sound separation. In speech enhancement Wiener filter can be used as time frequency mask where each mask gives relation between original speech signal and mixture of original and noisy signals. Wang and Brown [1] presented a monaural time-frequency masking algorithm. In this single microphone is used for the speech processing. Monaural algorithms uses sound properties for speech signal separation or auditory scene analysis. There are two types of algorithms: feature- based and model based. In both these algorithms binary or ratio (soft) mask is produced. Time frequency masking technique is based on two aspects: auditory masking and sparseness in the signals in T-F domain. P. Scalart and Filho [2] presented Wiener

algorithm depends on signal-to-noise ratio calculation and two spectral-subtractive algorithms based on reduced delay convolution [3]. In this algorithm, for sentence segmentation overlapping method is used with 160 samples and 50 % overlap is obtained. Hann window technique and discrete Fourier transform (DFT) is used for each segment. Performance of subtractive algorithms [7], the performance of Wiener algorithm is substantially higher. So, in hearing aid prosthetics Wiener algorithm is preferred than spectral-subtractive algorithm. J. Li, S. Sakamoto, S. Hongo, M. Akagi and Y. Suzuki [4] presented binaural algorithm for speech im-provement. The binaural algorithm is based on both CASA and ICA analysis and uses two or more microphones for speech separation. These algorithms only improves speech quality but no significant results are obtained for speech intelligibility. The proposed algorithm i.e. binary mask algorithm is suitable for real-time implementation of hearing prosthetics. V. Hanson [5] presented a binary mask algorithm that has a large potential for improving speech performance in background noise. In this algorithm, performance is effective even if the noise is not sparse in the time-frequency region. The whole SNR of the speech signal can be improved by discarding those regions of the time-frequency domain whose SNR is fails to go beyond a definite threshold level.

III. SYSTEM IMPLEMENTATION

Time-frequency masking is a widely used technique for speech and signal processing used in automatic speech recog- nition [5]. As shown in the lock diagram, mixture of two signals i.e. original speech and signal with noise acts as input for next stage. Mapping of speech from time to time frequency region is done. After that binary mask is created and with the help of SNR thresholding stage noise will be removed. Finally clean speech signal is obtained.

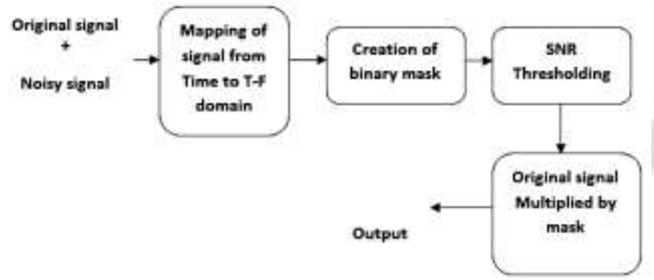

Fig. 2. Block diagram of binary mask algorithm.

A.Auditory Masking

As time frequency masking technique is based on scarcity of speech signals in T-F domain. Because of this, overlap- ping of noise with T-F domain region containing original speech signal is very difficult. Background noise which is very close to speech containing T-F regions that lowers the speech perceptibility by the psychoacoustic phenomenon i.e. auditory masking. Performance of speech intelligibility de- creases because of auditory masking even in normal SNR levels. Hearing impaired listeners find it more difficult due this auditory masking and loss of speech intelligibility. But those with normal hearing can handle this effect.

B. Improving Speech Intelligibility

Background noise is removed effectively by using binary mask algorithm. Noisy signals in the time frequency region masks original speech signal. Also it reduces speech clarity of signal. In this algorithm, in first step input signal in time domain is mapped T-F domain. The next step creates the binary mask by separating original speech signals from noisy signals. After this thresholding is done, recombined all the originals signal by removing noisy regions.

C. Binary Mask Algorithm Challenges

Spectral resolution is directly dependent on performance of binary mask algorithm. In spatial filtering stage, operations are performed sequentially but it requires higher clock rates. The chain of sequentially executed instructions are pipelined together. Memory requirement and hardware distributions will be reduced with proper selection of filter structure.

IV. EXPERIMENTAL ANALYSIS

A. Steps for Binary Mask Algorithm

1) Setting SNR: Noisy speech signals were generated in Matlab using a weighted sum of clean speech and noise. SNR levels were adjusted by changing the weight applied to the noise signal. SNR was measured using the global SNR output from the snrseg function in the Voice-box toolbox.

2) Audio record function : By using this function various speech samples are recorded with different background noise. These speech samples are real-time signals as it includes the target speech and background noise. The noise will be removed after applying binary mask algorithm.
3) Speech samples are recorded using single microphone only. Variables are initialised in the algorithm.
4) Allocate memory for various variables and noise mag- nitude is calculated.
5) Noise estimation is done for different parameters and divide real-time speech samples into frames.
6) Perform the short time Fourier transform (STFT) anal- ysis and compute true STFT noise spectrum.
7) Apply binary mask algorithm and create modified spec- trum and perform sanity check and reconstruct clean speech.
8) Apply inverse STFT and truncate the extra paddings used i.e. back to original signal length.

B. Implementation of Binary Mask algorithm

In this, speech samples are read from the audio database. Then the samples are stored in new variable in Matlab. Next step is to create noisy speech signal, for this noise is added to the clean speech signal(original speech signal). Create time vector for vector to frame conversion and vice versa. After this binary mask algorithm is applied to it. Output of the binary mask algorithm is speech samples which is written to the wave file for further playback. After listening the processed speech samples and noisy speech samples, conclusions are made that speech samples are enhanced effectively and background noise is removed significantly.

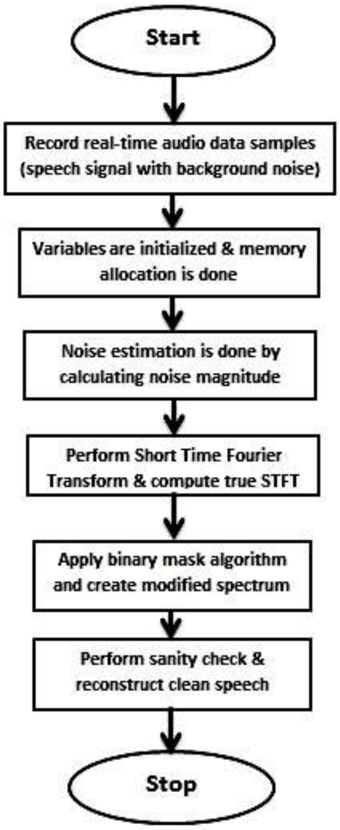

Fig. 3. Flow chart for Binary Mask algorithm.

V. RESULTS AND DISCUSSION

A. Simulation Results

The target signals used for testing the algorithm are taken from real-time recording. Also the target signals can be ob- tained from the TIMIT Acoustic Phonetic Continuous Speech Corpus. The specific noise chosen include babble (speech shaped noise), white noise and restaurant noise. Speech sim- ulation is done in Matlab. SNR values can be adjusted for various speech samples. SNR can be measured using global SNR output from snrseg function in voicebox toolbox MAT- LAB.

B. Algorithm demo using GUI

This demo provides a simple GUI for basic filtering of audio data. Using the GUI having advantages as following: 1. Load audio data stored in a .wav file. 2. Real-Time speech signals are recorded using binary mask algorithm. In this GUI, speech can be recorded or it can be loaded using stored database files. Add the filtered noised to the original audio signal. 4. Remove the noise by inverse filtering. 5. The purpose of this DEMO is not to provide robust and sophisticated denoising

algorithms, but simply to demonstrate some basic audio filtering processes in Matlab. In the above figure a screen shot of the demo is shown. In the first panel (row) you can see the plotted audio signal and the respective spectrogram. In the second panel

Fig. 4. Binary Mask Algorithm demo using GUI.

binary mask output waveform is shown and along with its spectrogram. In this GUI demo, real-time speech samples can be obtained by real-time recording. Also speech samples can be saved and can be used for future use. In the third panel, Power Spectral Density (PSD) for both real-time speech signal and binary mask output signal is shown respectively.

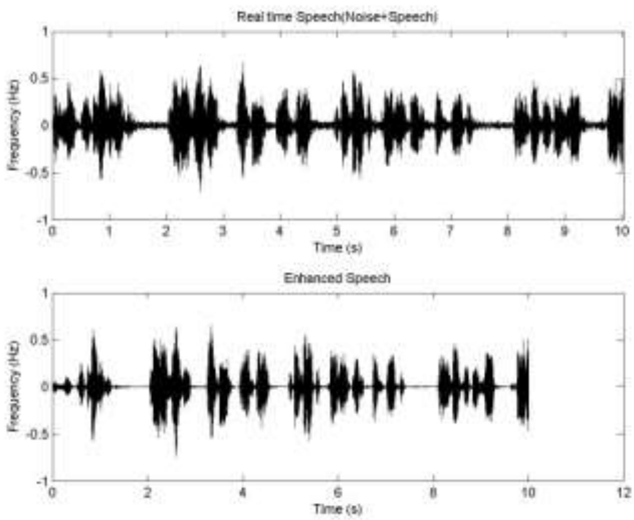

Fig. 5. Unprocessed and processed waveforms for sample 1.

In figure 7, the graph shows the percentage of improvement for binary mask algorithm over different speech enhancement algorithms. The speech samples are recorded real-time and for comparison 10 samples are taken.

VI. CONCLUSION AND FUTURE WORK

This proposed algorithm helps to enhance the clarity of speech signals having noise much higher as compared to other algorithms. Binary mask algorithm improves both speech quality as well as speech intelligibility. The proposed method is characterized by the following:

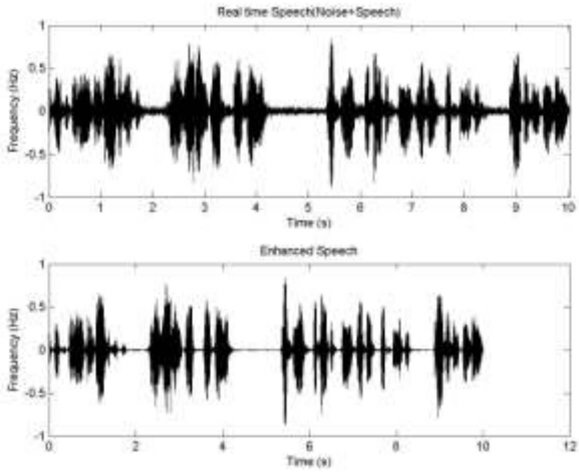

Fig. 6. Unprocessed and processed waveforms for sample

1)Implementation of bank of band-pass filters and binary mask algorithm improves as the frequency span of filter bank increases, which causes speech intelligibility to increase and noise sensitivity to decrease..

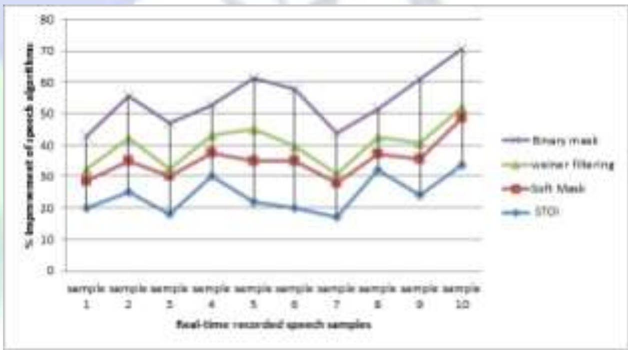

Fig. 7. Improvement of Binary Mask algorithm over other methods.

Filter bank is implemented in frequency range of 200 Hz to 10 kHz.

2) Filters are designed using Matlabs filter design tool.

3) For calculating binary mask, target+noise mixture and noise signals are performed in parallel using two separate banks of 28 band-pass filters.

4) Binary mask signal is created in clas- sification stage. Classification calculations are pipelined with spectral stage calculations to reduce the total processing time.

5) Target signals and noise signals are taken from databases. Future work can improve the proposed method by realizing on a low power DSP or in a low power ASIC. Binary mask algorithm can give give better results for speech intelligibility in low frequency range.

REFERENCES

[1] Wang D. L. and Brown G. J. ,"Computational auditory scene analysis: Principles, algorithms, and applications," Hoboken, NJ: Wiley/IEEE Press, 2006.

[2] P. Scalart and J. Filho, "Speech enhancement based on a priori signal to noise estimation," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., pp. 629632, 1996.

[3] H. Gustafsson, S. Nordholm, and I. Claesson,"Spectral subtraction using reduced delay convolution and adaptive averaging," IEEE Transaction on Audio processing., vol. 9, no.8, pp-799-807, Nov. 2001.

[4] J. Li, S. Sakamoto, S. Hongo, M. Akagi and Y. Suzuki, Two- Stage Bin- aural Speech Enhancement With Wiener Filter Based On Equalization-Cancellation Model, IEEE Workshop on Applications of Signal Process- ing to Audio and Acoustics, pp. 133 136 18-21, Oct. 2009.

[5] V. Hanson, K. Odame, "Real-Time Embedded Implementation of the Binary Mask Algorithm for Hearing Prosthetics," IEEE Transactions on Biomedical Circuits and Systems, vol. 8, No. 4, August 2014.

[6] V. Hanson and K. Odame, "Real-time source separation on a field programmable gate array platform," in Proc. Annual International Con- ference of IEEE Engineering in Medicine and Biology Soc., Aug. 2012, pp. 29252928.

[7] P. C. Loizou and G. Kim, "Reasons why current speech-enhancement al- gorithms do not improve speech intelligibility and suggested solutions," IEEE Transaction on Audio, Speech, Language Process., vol. 19, no. 1, pp.4756, Jan. 2011.

[8] C. H. Taal, R. C. Hendriks, R. Heusdens, and J. Jensen, "An algorithm for intelligibility prediction of time-frequency weighted noisy speech," IEEE Transaction on Audio, Speech, Language Process., vol. 19, no. 7, pp. 21252136, Sep. 2011.

[9] N. Li and P. C. Loizou, "Factors influencing intelligibility of ideal binary-masked speech: Implications for noise reduction," J. Acoust. Soc. Amer., vol. 123, no. 3, pp. 16731682, Mar. 2008.

[10] J. Boldt, U. Kjems, M. S. Pedersen, T. Lunner, and D. Wang, "Estimation of the ideal binary mask using directional systems," in Proc.11th Int. Workshop Acoustic Echo and Noise Control, pp.395401, Sep. 2008.

[11] N. Marrone, C. R. Mason, and G. Kidd, Jr., "The effects of hearing loss and age on the benefit of spatial separation between multiple talkers in reverberant rooms," J. Acoust. Soc. Amer., vol. 124, no. 5, pp. 30643075, Nov. 2008.

[12] D.Wang, "Time-frequency masking for speech separation and its po- tential for hearing aid design," Trends Amplification, vol. 12, no. 4, pp.332353, Dec. 2008.

[13] D. Wang, U. Kjems, M. Pedersen, J. Boldt and T. Lunner, "Speech intelligibility in background noise with ideal binary time-frequency masking," J. Acoustical Society of America, vol. 125, no. 4. pp. 23362347, April 2009.

[14] A. V. Oppenheim et al., "Discrete-Time Signal Processing". Upper Saddle River, NJ, USA: Prentice-Hall, vol. 5, 1999.

[15] G. W. Elko and A.-T. N. Pong, "A steerable and variable first-order differential microphone array," in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing, , pp. 223226, Apr. 1997.

[16] G.W. Elko and A.-T. N. Pong, "A simple adaptive first-order differential microphone," in Proc. IEEE ASSP Workshop Applications of Signal Processing to Audio and Acoustics, , pp. 169172, Oct. 1995.

[17] DeLiang Wang, ”Time-frequency Masking for Speech Separation and Its Potential for Hearing Aid Design”, Trends in Amplification, Oct. 2008.

AUTHOR’S BIBLOGRAPHY

Swati Pawar received the B.Tech degree in Electronics and Telecommunication from Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, in 2012 and currently pursuing masters degree from Sinhgad Institute, Savitribai Phule University, Pune. She had published research paper related masters degree project work in IEEE Transaction. Her research interest includes embedded systems, android, micro controllers, reconfigurable computing architectures and audio and video processing.

